PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, C07K 14/435, G01N 33/574, C12Q 1/68

A1

(11) International Publication Number: **WO 96/37611**

(43) International Publication Date: 28 November 1996 (28.11.96)

(21) International Application Number: PCT/US96/07747

(22) International Filing Date: 22 May 1996 (22.05.96)

(30) Priority Data:
08/446,919 22 May 1995 (22.05.95) US

(71) Applicant: THE JOHNS HOPKINS UNIVERSITY [US/US]; 720 Rutland Avenue, Baltimore, MD 21205 (US).

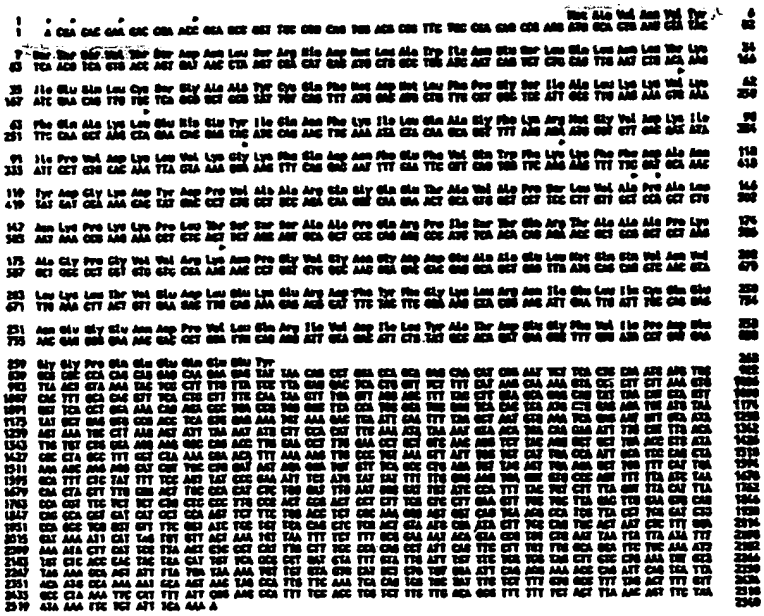
(72) Inventors: VOGELSTEIN, Bert; 3700 Breton Way, Baltimore, MD 21208 (US). KINZLER, Kenneth; 1348 Halstead Road, Baltimore, MD 21234 (US).

(74) Agents: KAGAN, Sarah, A. et al.; Banner & Allegretti, Ltd., 11th floor, 1001 G Street, N.W., Washington, DC 20001-4597 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: EB1 GENE PRODUCT BINDS TO APC



(57) Abstract

Inactivation of the APC tumor suppressor gene plays an important role in the development of both sporadic and familial forms of colorectal cancers. The majority of these mutations result in the loss of the carboxyl terminus of the APC protein. A cellular protein, EB1, that associates with the carboxyl terminus of APC both *in vitro* and *in vivo* has been identified. The EB1 gene is predicted to encode a 268 amino acid protein without significant homology to any protein with known function.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Latvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

EB1 GENE PRODUCT BINDS TO APC

This invention was made with support from the National Institutes of Health, Grant No. CA57345. The U.S. government therefore retains certain rights in the invention.

BACKGROUND OF THE INVENTION

The *APC* gene (adenomatous polyposis coli) was originally isolated by virtue of its alteration in familial and sporadic forms of colorectal cancer (1-4). Germline mutations of the *APC* gene account for most cases of familial adenomatous polyposis (FAP), an autosomal, dominantly inherited disease that predisposes patients to multiple colorectal polyps and cancer (reviewed in 5). APC mutations have also been found in cancers of the central nervous system. While FAP patients with germline mutations of *APC* account for less than 1% of colorectal cancers in the United States, somatic mutations of *APC* occur in the majority of colorectal adenomas and cancers (6-9). These alterations appear to occur early as they can be identified in the smallest identifiable lesions including dysplastic aberrant crypt foci (6,10,11). The vast majority of both germline and somatic *APC* mutations are predicted to result in truncation of the APC protein due to either nonsense or frame-shifting mutations (5,6,7,8,9). Likewise, mice carrying homologous germline truncating mutations of *APC* are also predisposed to intestinal tumors (8, 9, 10). Altogether, these results strongly suggest that *APC* mutations are an early if not initiating event in the development of both sporadic an inherited forms of colorectal cancer.

While disruption of normal APC function clearly plays a role in colorectal tumorigenesis, what this function might be remains unclear. The *APC* gene is

predicted to encode a protein of 2843 amino acids with limited functional homology to known proteins. The primary structure contains several Armadillo repeats that are shared by proteins with apparently diverse functions (3, 15) as well as several regions of heptad repeats of the type that mediate oligomerization via coiled-coil structures (3). Indeed, the amino terminus of APC, which has a very strong potential for forming coiled-coil structures, has been shown to mediate the homo-oligomerization of APC protein (16, 17). Three additional repeats located between amino acids 1000 and 1200 of APC mediate an associate with α and β - catenins, critical cytoplasmic components of cadherin cell adhesion (18, 19). In addition, wild-type but not mutant forms of APC have been shown to associate with microtubule cytoskeleton (20, 21).

While the aforementioned biochemical characteristics of APC provide important clues to its function, other functions remain undefined. Because mutant APC proteins almost uniformly lack their carboxyl terminus, we hypothesized that the carboxyl terminus of APC interacts with proteins that are essential for its normal function. To test this hypothesis we attempted to identify a protein that associates with the carboxyl terminus of APC.

SUMMARY OF THE INVENTION

It is an object of the invention to provide a nucleic acid molecule encoding a protein which binds to APC.

It is an object of the invention to provide a protein molecule which binds to APC.

It is another object of the invention to provide nucleic acid molecules which can be used to detect genes involved in neoplasia in a sample.

It is yet another object of the invention to provide methods for determining a predisposition to colorectal and other neoplasms.

It is still another object of the invention to provide antibodies useful for detecting proteins which bind to APC.

It is an object of the invention to provide methods for assessing susceptibility to colorectal and other cancer.

It is an object of the invention to provide methods for diagnosing cancer.

It is still another object of the invention to provide methods to assess treatment options for a cancer.

It is yet another object of the invention to provide methods to assess the status of *APC* alleles in a cell.

These and other objects of the invention are provided by one or more of the embodiments described below. In one embodiment of the invention a nucleic acid molecule is provided which comprises an *EB1* DNA according to SEQ ID NO:1. Also provided is a molecule which may contain at least 12, 18, or 20 contiguous nucleotides of *EB1* coding sequence. Also provided is a molecule which encodes at least about 6, 8, 10, or 20 contiguous *EB1* amino acids.

In another embodiment of the invention an isolated and purified EB1 protein is provided. The protein has an amino acid sequence according to SEQ ID NO:2. Polypeptides having at least 6, 8, 10, or 20 contiguous amino acids of said sequence are also provided.

In still another embodiment of the invention a method for determining a predisposition to or a diagnosis of colorectal and other neoplasms is provided. The method comprises the step of: determining one or more mutations in one or more *EB1* alleles of a human tissue, wherein wild-type *EB1* is as shown in SEQ ID NO:1.

In one embodiment of the invention a method for determining a predisposition to or diagnosis of colorectal and other neoplasms is provided. The method comprises the step of: assaying protein complexes in a cell, wherein said protein complexes comprise APC and EB1, wherein absence of said complexes or reduction in level of said complexes indicates a predisposition to neoplasms.

In another embodiment of the invention an antibody preparation is provided. The antibody is specifically immunoreactive with an EB1 protein according to SEQ ID NO:2.

According to still another aspect of the invention a method for determining a diagnosis or predisposition to cancer is provided. The method comprises the

step f: testing a human tissue to determine if the tissue expresses less EB1 gene product than a normal human tissue or no EB1 gene product.

In another embodiment of the invention a method is provided to assess treatment options for a cancer. The method comprises the step of: contacting a lysate of cancer cells with EB1 protein and detecting the formation of protein complexes comprising said EB1 protein, a lysate which fails to form complexes indicating cancer cells which are good candidates for treatment with cyclooxygenase inhibitors.

In yet another embodiment of the invention, a method is provided to assess the status of *APC* alleles in a cell. The method comprises the step of contacting a lysate of cells with EB1 protein, a lysate which fails to form complexes indicating cancer cells which may lack wild-type *APC*.

These and other embodiments of the invention provide the art with the identity of a gene and a protein which are involved in the suppression of neoplasia.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the nucleotide and predicted amino acid sequences of EB1. The arrowheads above the sequences indicate the 5' termini of different *EB1* cDNA clones isolated by yeast two hybrid screening. The predicted amino acid sequence begins at nucleotide 65 and ends at the nucleotide 868. The nucleotide sequence has been deposited with Genbank (# U24166).

Figure 2 shows *in vitro* Binding of EB1 to APC. Figure 2A shows binding of cellular APC to GST-EB1 (glutathione S-transferase = GST) fusion protein. SW480 and HCT116 are human colorectal cancer cell lines that express truncated and full length APC, respectively (19). Protein from total cell lysates (-) or protein bound by GST-EB1 fusion protein (GST-EB1) were analyzed by Western blot analysis with APC-specific monoclonal antibody FE9 (19). Figure 2B shows the binding of EB1 to GST-APC fusion protein. GST-CTN has been described (19) and was used as a negative control. SW480 and HCT116 cells were metabolically labelled with ^{35}S-Met and incubated with the GST fusion proteins as indicated. *In vitro* transcribed and translated EB1 (in vitro) was run on gel

directly (-) or following binding to GST-APC(X) fusion protein as indicated. Proteins were detected by fluorography. Figure 2C shows one dimensional peptide mapping. Cellular (SW480, HCT116) and *in vitro* translated (in vitro) EB1 proteins were isolated by binding to GST-APC(X) and subjected to one dimensional peptide mapping as described (19).

Figure 3 shows *in vivo* association of APC and EB1. SW480 cells were transiently transfected with expression vectors for EB1 or APC as indicated. The parental expression vector pCMV-NEO-BAM (pCMV) was used to equalize the total amount of DNA transfected. Lysates prepared from these transfected cells were used directly (total), or after immunoprecipitation with a monoclonal antibody against hemagglutinin (HA) as negative control or an EB1-specific monoclonal antibody (EB1). Detection of APC was carried out by immunoblotting using APC specific monoclonal antibody FE9. MT and FL indicate truncated and full length APC, respectively.

Figure 4 shows the localization of EB1 to chromosome 20q11.2 by fluorescence in situ hybridization (FISH). The left panel shows an ideogram of a G-banded human chromosome 20 with the band q11.2 bracketed. The top right panel shows the fluorescent signals localizing *EB1* to chromosome 20. The bottom right panel shows a G-banded human chromosome 20 localizing *EB1* to 20q11.2.

Figure 5 shows human and yeast EB1 homologues. Figure 5A shows an amino acid sequence comparison among human EB1 homologues. EB2 represents the amino acid sequence predicted from the nucleotide sequence of a contig of 3 different EST's (Z46175, T17004 and Z42534.) The Z19434 and M85402 lines show the predicted amino acid sequences of these two EST's, respectively. Because of the lack of overlap between Z19434 and M85402, we could not determine whether they represented one or two genes. "-" indicates that no sequence information was available at that position. Figure 5B shows an amino acid sequence comparison between human EB1 and a potential yeast EB1 homolog. The sequence of Yer016p is predicted from an open reading frame (ORF) from yeast chromosome V as described in the text. "-" indicates gap introduced to

allow the best alignment between the two sequences. In both Figures 5A & 5B, blocks of homology are capitalized and shaded according to their mean scores.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

We have identified a cDNA that codes for a protein that interacts with the carboxyl terminus of APC. This interaction was clearly demonstrated by binding of cellular APC to recombinant EB1 and by binding of cellular EB1 to recombinant APC. The association between EB1 and APC in mammalian cells was also demonstrated in cells cotransfected with vectors expressing these two proteins. Because almost all previously identified *APC* mutations result in the truncation of the APC protein, these mutant APC proteins cannot associate with EB1. This observation strongly suggests that the interaction between APC and EB1 is important for the normal function of APC and that loss of this association is essential for the development of colorectal cancer. Mutation of EB1 is one way that a cell can lose this association.

EB1 nucleic acid molecules according to the present invention include both ribonucleic acids and deoxyribonucleic acids. They may be incorporated as a part of a vector, such as a virus, phage, plasmid, minichromosome, etc. A vector typically contains an origin of replication which allows for independent replication of the nucleic acids of the vector and any insert it may be carrying. Suitable vectors may be chosen for a particular purpose, as is well within the skill of the art. Isolation and purification of nucleic acid molecules from other nucleic acid molecules and from other cellular components can be accomplished as is well known in the art. Nucleic acid molecules comprising at least about 12, 18, or 20 nucleotides of *EB1* coding sequence can be used *inter alia* as probes and primers. Probes are typically labelled with a detectable label such as a radionuclide, an enzyme, or ligand. Primers may have restriction enzyme sites or promoters appended, as may be desirable for cloning or *in vitro* protein synthesis. Nucleic acid molecules encoding at least about 6 or 20 contiguous amino acids of *EB1* can be used for expressing fragments of EB1, for example for use in fusion proteins or as antigens or immunogens. The nucleotide sequence of wild-type *EB1* is

provided in SEQ ID NO:1. The amino acid sequence f EB1 protein is provided in SEQ ID NO:2.

EB1 protein may be isolated and purified from human cells, from transformed mammalian, other eukaryotic, or prokaryotic cells. Purification may be accomplished employing antibodies which are specific for EB1, such as AE9, EA3, and GD10, as provided herein. Other antibodies can be used which are made using all or a portion of EB1 as an immunogen. Affinity methods may also be used which take advantage of the binding of EB1 to APC. EB1 may also be synthesized chemically or in an *in vitro* system, as described in more detail below. Portions of EB1 which contain at least 6 or 20 contiguous amino acids according to SEQ ID NO:2 can be used in assays and as immunogens. These can be synthesized and isolated according to established techniques with the benefit of the sequence information provided herein.

Predisposition to colorectal and other neoplasms can be determined by examination of a sample for a mutation in an *EB1* gene. Such other cancers include, but are not limited to desmoid tumors, osteomas, glioblastomas, medulloblastomas and other tumors of the central nervous system. Examination can be done by comparison with the wild-type sequence provided in SEQ ID NO:1 or to the *EB1* found in human tissues which are normal. It can also be done by determining diminished expression of EB1 protein or message, or failure of EB1 to form complexes with APC. Methods for determining mutations include PCR, sequencing, restriction mapping, S1 nuclease mapping, and hybridization with allele-specific probes. Any method known in the art can be used. Methods for determining diminished EB1 expression or failure to form complexes with APC can be determined using techniques such as immunoprecipitation, immunoblotting, immunohistochemistry, etc. Antibodies which are particularly useful for such purposes are monoclonal antibodies AE9, EA3 and GD10, whose isolation and production are discussed in more detail below. Polyclonal antibodies can also be used, especially if purified to render a preparation monospecific. Samples which may be tested for assessing susceptibility to colorectal cancer include blood,

chorionic villi, fetal trophoblasts, amniotic fluid, and blastomeres of pre-implantation embryos. Solid tissues can also be tested to determine predisposition and/or diagnosis.

Assays using EB1 can be used to assess the status of *APC* alleles, since according to the present invention EB1 and APC interact. Thus, for example, a lysate of cells can be contacted with EB1 protein and the formation of protein complexes comprising EB1 protein can be detected. If the lysate fails to form complexes with EB1 the cells are likely cancer cells which lack wild-type *APC*. Other means for measuring the interaction of EB1 with APC can be used to provide such information.

The drug sulindac has been shown to inhibit the growth of benign colon tumors in patients with familial adenomatous polyposis (FAP), presumably by virtue of its activity as a cyclooxygenase inhibitor (Waddell et al., J. Surg. Oncology 24(1), 83 (1983); Wadell, et al., Am. J. Surg., 157(1), 175 (1989); Charneau et al., Gastroenterologie Clinique at Biologique 14(2), 153 (1990)). Cyclooxygenase is required to convert arachidonic acid to prostaglandins and other biologically active molecules. Since FAP is attributed to mutations in *APC*, treatment options for a cancer may be assessed using EB1. EB1 can be used as described above to assess the status of *APC* alleles. Cells which fail to form protein complexes with EB1 are likely cancer cells which are good candidates for treatment with cyclooxygenase inhibitors, such as sulindac.

EXAMPLES

Example 1

This example describes the isolation of a gene which encodes a protein which interacts with the carboxy terminus of APC.

We used a modified yeast two hybrid system (22,23) to screen a HeLa cDNA library for proteins interacting with the carboxyl terminus (codons 2167 to 2843) of APC. A total of 90 positive clones with the appropriate phenotype were identified after screening one million transformants. The cDNAs isolated from 67

out of these 90 clones were abl to confer the correct phenotype when retransformed into the test strain of yeast. The nucleotide sequences of both ends of each cDNA were determined and were compared to each other. Forty-eight of these cDNAs were found to be derived from a same gene and could be separated into 11 groups according to their length (Figure 1). We chose to characterize this cDNA in detail and named it EB1 (for EcoRI fragment binding protein 1). The fusion proteins encoded by two independent cDNA clones did not interact with amino proximal residues 6 to 1013 when tested in the two hybrid assay.

Northern blot analysis with probes to EB1 identified a single 2.4 kb transcript. Because the largest EB1 cDNA isolated by interaction trap method was 1.4 kb, we screened a human fetal brain cDNA library to isolate the full length cDNA. None of the newly isolated cDNA clones had additional 5' nucleotide sequence but many of them had additional 3' nucleotide sequence extending the length of the cloned message to 2.4 kb. Furthermore, no additional 5' sequence was obtained after screening three 5'-RACE cDNA libraries. Together, these results suggest that the full-length message for EB1 had been isolated. Nucleotide sequence analysis of the overlapping cDNA clones revealed an ORF extending from nucleotide 1 to 868 (Figure 1). If translation initiated at the first methionine, *EB1* would be predicted to encode a 268 amino acid protein with a predicted molecular weight of 30 kD.

Methods: Two hybrid screening. The modified yeast two hybrid system, the cDNA library and screening the cDNA library using this system have been described (22, 23). The bait was made by inserting a 2.5 Kb EcoRI fragment of *APC* containing nucleotides nucleotide 6498 to 8950 into the SmaI site of LexA(1-202)+PL (24) after making the EcoRI fragment blunt-ended using the Klenow fragment of DNA polymerase I.

Example 2

This example demonstrates the *in vitro* and *in vivo* binding of APC to EB1.

To confirm and extend the two hybrid results, we tested the direct interaction between EB1 and APC using an *in vitro* binding assay. The carboxyl

terminal 163 residues of EB1 were expressed as a glutathione-S-transferase fusion protein in *E. coli*. This fragment was expected to bind APC because it included more of EB1 than several of the *EB1* cDNA clones originally isolated by the yeast interaction trap method. As expected, this fusion protein was able to associate with the full-length APC from cell lysates, but was unable to bind to mutant APC that lacked the putative EB1 binding region (Figure 2A). This result clearly showed that EB1 interacts with endogenous APC and that this interaction requires the carboxyl terminus of APC.

To test whether APC could bind endogenous EB1, we expressed amino acid codons 2167 to 2843 of APC as a GST fusion protein (GST-APCE) and incubated the purified fusion protein bound on the glutathione agarose with lysates prepared from metabolically labeled colon cancer cell lines. The APC fusion protein bound a 30 kD cellular protein bound which had identical mobility to the EB1 expressed *in vitro* (Figure 2B). To confirm that this 30 kD protein was indeed EB1, we compared the one-dimensional peptide map of this 30 kD protein with that of EB1 expressed *in vitro*. The peptide maps of these proteins were identical (Figure 2C). This result also provided additional evidence that the first codon for methionine in the EB1 cDNA is the translational initiation codon.

Methods: GST fusion proteins. The pGSTagEB1A expression vector was constructed using an EcoRI fragment (nucleotides 317 to 899 of EB1) of an EB1 cDNA clone isolated by interaction trap screening. After subcloning into the EcoRI site of pBluescript SK II, the EcoR1 fragment was excised as a BamHI-SalI fragment and inserted into the BamHI and XhoI sites of pGSTag (25). The pGSTagEB1B expression vector constructed by inserting a 1.8 Kb SalI-HindIII fragment (nucleotides 40 to 2091) of an EB1 cDNA clone isolated from human fetal brain cDNA library into the SalI and HindIII sites of pGSTag. The pGSTagAPCE expression vector was constructed by inserting the 2.5 Kb EcoRI fragment of *APC* cDNA, identical to that used for making the bait for two hybrid screening, into the EcoRI site of pGSTag. The expression and purification of fusion proteins were carried out as described (19).

Methods: PCR and *in vitro* expression of EB1. The EB1 coding region was amplified by using the upstream primer 5'-GGATCCTAATACGACTCACTATAGGGAGACCACCATGGCAGTGAACG TATACTC-3' and the downstream primer 5'-ATTTCTCCACTGAGGTCGC-3'. The upstream primer contains the sequence of the promoter for the T7 DNA polymerase and the first 20 nucleotides of the EB1 coding sequence. The downstream primer locates at the 3' untranslated region of EB1. The PCR reaction was carried out using an isolated cDNA clone as the template with 35 cycles of 30 sec at 95°C, 1 min at 50°C, and 1 min at 70°C. The PCR product was using directly in a coupled *in vitro* transcription and translation reaction as described (26).

Methods: *in vitro* binding assay. Metabolically labelled protein extracts from the human colorectal cancer cell lines SW480 and HCT116 were used for the *in vitro* binding assay. Metabolic labeling, preparation of cell lysates, *in vitro* binding, and peptide mapping were carried out as described (19).

Example 3

This example demonstrates the *in vivo* association of EB1 and APC by co-immunoprecipitation.

In order to further characterize the association APC and EB1, three monoclonal antibodies (AE9, EA3 and GD10) against EB1 were generated. Western blot analysis with all three of these antibodies detected a 30 kD protein in total cell lysates which associated with GST-APCE, but not with a control protein GST-CTN. EB1 protein was detected in several human colon cancer cell lines including a human kidney fibroblast cell line 293, the canine kidney epithelial cell line MDCK, and the mouse fibroblast cell line NIH3T3. To demonstrate an *in vivo* association between EB1 and APC in mammalian cells, SW480 cells were transiently transfected with vectors expressing APC or EB1. The association between these two proteins was examined by immunoprecipitation using the EB1-specific antibody EA3 followed by immunoblotting with the APC-specific antibody FE9. The co-immunoprecipitation of APC and EB1 was clearly demonstrated

when cells were transfected with both expression vectors but not when either one was omitted. (Figure 3.)

We have not been able to detect the association between endogenous full-length APC and EB1 by co-immunoprecipitation experiments. The reason for this may be purely technical. This is consistent with our inability to co-immunoprecipitate APC and EB1 from cell lysates prepared from yeast clones with clear functional evidence of an association between these two proteins as reflected by the two-hybrid assay. Similar reasons have also been suggested for the failure to demonstrate an association between pRB and RBP2 by co-immunoprecipitation (26, 27).

Methods: Monoclonal antibodies. The three EB1 monoclonal antibodies, AE9, EA3, and GD10, were derived from mice immunized with GST-EB1 fusion protein. Immunization of mice, cell fusion, and the preparation of monoclonal antibodies were carried out as described (27). The EA3 monoclonal was found to specifically recognize EB1 by both Western blot and immunoprecipitation.

Methods: *in vivo* Binding Assay. SW480 cell lines were transiently transfected with pCMV-APC or pCMV-EB1. The pCMV-APC was as described (20) and the pCMV-EB1 vector was derived by cloning a PCR product containing EB1 nucleotides 62 to 871 into the BamH1 site of pCMV-NEO-BAM. PCR was performed with following primers which were engineered to include the underlined BglII sites: 5'-CGAGATCTAAGATGGCAGTGAACGTATAC-3' and 5'-GCAGATCTTTAATACTCTTCTTGATCCTCC-3'). To eliminate the possibility of PCR errors, the sequence of the *EB1* fragment cloned into PCMV-EB1 was verified by nucleotide sequencing. Transient transfections, preparation of cell lysates, immunoprecipitation and western blot analysis were performed as described (16, 19, 20).

Example 4

This example demonstrates the chromosomal mapping of EB1.

The chromosomal localization of EB1 was determined by fluorescence in situ hybridization (FISH). Three P1 clones for EB1 were isolated from a P1

library by PCR. One of these P1 clones was used as the probe in the FISH analysis as previously described (24). Sixteen out of a total of 17 metaphase cell examined displayed double fluorescent signals (i.e. one on each chromatin) on the proximal short arm of chromosome 20. The same cells hybridized for FISH had been previously G-banded and photographed to allow direct comparisons of the results. The result demonstrated that the sequences hybridizing to EB1 can be localized to 20q11.2 (Figure 4).

Methods: Chromosomal localization. Three EB1 genomic clones (EB-922, EB1-923, EB1-924) were obtained by PCR screening of A P1 library (Genome Systems, Inc.) using primers (5'-AAAACAGAGAGGCTGACCG-3 and 5'-ATTTCTCCACTGAGGTCGC-3') designed to amplify EB1 nucleotides 1102 to 1205. Total EB1-923 DNA was labeled with Biotin-16-dUTP by nick translation and used for FISH. For FISH, about 100 ng of probe was used in 10 μl hybridization mixture (55% formamide, 2X SSC, and 1 μg human Cot 1 DNA) which was denatured at 75°C for 5 minutes. Hybridization was carried out using a modified procedure of Pinkel et al. (28) as previously described (29).

Example 5

This example analyzes the nucleotide and amino acid sequences of EB1.

Searches of the National Center for Biotechnology Information (NCBI) non-redundant nucleotide and EST (expressed sequence tag) databases indicated that EB1 had not been previously characterized although there were several ESTs that were almost identical to parts of the 3' untranslated region. Interestingly, there were also five ESTs which were similar but not identical to the coding region of EB1. These ESTs likely represented novel EB1-related genes rather than sequencing mistakes as there were numerous nucleotide substitutions that preserved the encoded amino acids of *EB1* in these ESTs. These five ESTs could be divided into three contigs which represented at least two different EB1 related proteins (Figure 5A). Searches of NCBI's non-redundant protein database with EB1 identified three proteins with statistically significant ($P < 0.05$) multiple regions of homology. These were a calcium channel protein from carp (PIR# A37860, P

= .0075), a bacterial RNA polymerase sigma chain homolog (PIR # JN0445, P = .0028) and Yer016p ($P = 2.4 \times 10^{-53}$). Yer016p is a putative gene identified in a 66,030 bp *Saccharomyces cerevisiae* chromosome V cosmid contig (Genbank #U18778). The predicted Yer016p protein shared five blocks of similarity with EB1 and could represent a yeast homolog of EB1 (Figure 5B). Together, these data suggest that EB1 is a member of a highly conserved multi-gene family.

Methods: Database searches and alignments. The NCBI's non-redundant nucleotide, non-redundant protein and DBEST databases (1/19/95 releases) were searched using the BLASTN, BLASTP and TBLASTN basic local alignment search software, respectively (30). Multiple alignments were performed using the MACAW multiple alignment construction and analysis software version 2.03 (31).

References

1. Groden, J., Thliveris, A., Samowitz, W., Calson, M., Gelbert, L., Albertsen, H., Joslyn, G., Stevens, J., Spirio, L., Robertson, M., Sargeant, L., Krapcho, K., Wolff, E., Burt, R., Hughes, J. P., Warrington, J., McPherson, J., Wasmuth, J., Le Paslier, D., Abderrahim, H., Cohen, D., Leppert, M., and White, R. Identification and characterization of the familial adenomatous polyposis coli gene. *Cell*, *66*: 589-600, 1991.
2. Joslyn, G., Calson, M., Thliveris, A., Albertsen, H., Gelbert, L., Samowitz, W., Groden, J., Stevens, J., Spirio, L., Robertson, M., Sargeant, L., Krapcho, K., Wolff, E., Burt, R., Hughes, J. P., Warrington, J., McPherson, J., Wasmuth, J., Le Paslier, D., Abderrahim, H., Cohen, D., Leppert, M., and White, R. Identification of deletion mutations and three new genes at the Familial polyposis locus. *Cell*, *66*: 601-613, 1991.
3. Kinzler, K. W., Nilbert, M. C., Su, L.-K., Vogelstein, B., Bryan, T. M., Levy, D. B., Smith, K. J., Preisinger, A. C., Hedge, P., McKechnie, D., Finniear, R., Markham, A., Groffen, J., Boguski, M. S., Altschul, S. F., Horii, A., Ando, H., Miyoshi, Y., Miki, Y., Nishisho, I., Nakamura, Y. Identification of FAP locus genes from chromosome 5q21. *Science*, *253*: 661-665, 1991.
4. Nishisho, I., Nakamura, Y., Miyoshi, Y., Miki, Y., Ando, H., Horii, A., Koyama, K., Utsunomiya, J., Baba, S., Hedge, P., Markham, A., Kruch, A. J., Petersen, G., Hamilton, S. R., Nilbert, M. C., Levy, D. B., Bryan, T. M., Preisinger, A. C., Smith, K. J., Su, L.-K., Kinzler, K. W.,

Vogelstein, B. Mutations of chromosome 5q21 genes in FAP and col rectal cancer patients. *Science*, *253*: 665-669, 1991.

5. Nagase, H., and Nakamura, Y. Mutations of the *APC* (adenomatous polyposis coli) gene. *Human Mutation 2:* 425-434, 1993.
6. Powell, S. M., Zilz, N., Beazer-Barclay, Y., Bryan, T. M., Hamilton, S. R., Thibodeau, S. N., Vogelstein, B., and Kinzler, K. W. *APC* mutations occur early during colorectal tumorigenesis. *Nature*, *359*: 235-237, 1992.
7. Miyoshi, Y., Nagase, H., Ando, H., Horii, A., Ichii, S., Nakatsuru, S., Aoki, T., Miki, Y., Mori, T., and Nakamura, Y. Somatic mutations of the *APC* gene in colorectal tumors: mutation cluster region in the *APC* gene. *Human Molecular Genetics*, *1*: 229-233, 1992.
8. Miyaki, M., Konishi, M., Kikuchi-Yanoshita, R., Enomoto, M., Igari, T., Tanaka, K., Muraoka, M., Takahashi, H., Amada, Y., Fukayama, M., Maeda, Y., Iwama, T., Mishima, Y., Mori, T., and Koike, M. Characteristics of somatic mutation of the adenomatous polyposis coli gene in colorectal tumors. *Cancer Research*, *54*: 3011-3020, 1994.
9. De Benedetti, L., Sciallero, S., Gismondi, V., James, R., Bafico, A., Biticchi, R., Masetti, E., Bonelli, L., Heouaine, A., Picasso, M., Groden, J., Robertson, M., Risio, M., Caprilli, R., Bruzzi, P., White, R..L., Aste, H., Santi, L., Varesco, L., and Ferrara, G.B. Association of *APC* gene mutations and histological characteristics of colorectal adenomas. *Cancer Research*, *54*: 3553-3556, 1994.
10. Jen, J., Powell, S.M., Papadopoulos, N., Smith, K.J., Hamilton, S.R., Vogelstein, B., and Kinzler, K.W. Molecular determinants of dysplasia in colorectal lesions. *Cancer Research*, *54*: 5523-5526, 1994.
11. Smith, A.J., Stern, H.S., Penner, M., Hay, K., Mitri, A., Bapat, B.V., and Gallinger, S. Somatic *APC* and K-ras codon 12 mutations in aberrant crypt foci from human colons. *Cancer Research*, *54*: 5527-5530, 1994.
12 Moser, A. R., Pitot, H. C., and Dove, W. F. A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse. *Science*, *247*: 322-324, 1990.
13. Su, L.-K., Kinzler, K. W., Volgelstein, B., Preisinger, A. C., Moser, A. R., Luongo, C., Gould, K. A., and Dove, W. F. Multiple intestinal neoplasia caused by a mutation in the murine homolog of the *APC* gene. *Science*, *256*: 668-670, 1992.
14. Fodde, R., Edelmann, W., Yang, K., van Leeuwen, C., Carlson, C., Renault, B., Breukel, C., Alt, E., Lipkin, M., Meera Khan, P., and Kucherlapati, R. A targeted chain-termination mutation in the mouse *Apc* gene results in mutiple intestinal tumors. *Proc. Natl. Acad. Sci. USA*, *91:* 8969-8973, 1994.
15. Peifer, M., Berg, S., and Reynolds, A. B. A repeating amino acid motif shared by proteins with diverse cellular roles. *Cell 76:* 789-791, 1994.

16. Su, L.-K., Johnson, K. A., Smith K. J., Hill, D. E., Vogelstein, B., and Kinzler, K. W. Association between wild-type and mutant *APC* gene products. *Cancer Research, 53:* 2728-2731, 1993.
17. Joslyn, G., Richardson, D. S., White, R., and Alber, T. Dimer formation by an N-terminal coiled-coil in the APC protein. *Proc. Natl. Acad. Sci. USA, 90:* 11109-11113, 1993.
18. Rubinfeld, B., Souza, B., Albert, I., Muller, O., Chamberlain, S. H., Masiarz, F. R., Munemitsu, S., and Polakis, P. Association of the *APC* gene product with beta-catenin. *Science, 262:* 1731-1733, 1993.
19. Su, L.-K., Vogelstein, B., and Kinzler, K. W. Association of the APC tumor suppressor protein with catenins. *Science, 262:* 1734-1737, 1993.
20. Smith, K. J., Levy, D. B., Maupin, P., Pollard, T. D., Vogelstein, B., and Kinzler, K. W. Wild-type but not mutant *APC* associates with the microtubule cytoskeleton. *Cancer Research, 54:* 3672-3675, 1994.
21. Munemitsu, S., Souza, B., Muller, O., Albert, I., Rubinfeld, B., and Polakis, P. The *APC* gene product associates with microtubules *in vivo* and promotes their assembly *in vitro*. *Cancer Research, 54:* 3676-3681, 1994.
22. Fields, S., and Song, O.-K. A novel genetic system to detect protein-protein interactions. *Nature, 340:* 245-246, 1989.
23. Gyuris, J., Golemis, E., Chertkov, H., and Brent, R. Cdi1, a human G1 and S phase protein phosphatase that associates with cdk2. *Cell, 75:* 791-803, 1993.
24. Ruden, D. M., Ma, J., Li, Y., Wood, K., and Ptashne, M. Generating yeast transcriptional activators containing no yeast protein sequences. *Nature, 350:* 250-252, 1991.
25. Ron, D., and Dressler, H. pGSTag--A versatile bacterial expression plasmid for enzymatic labeling of recombinant proteins. *BioTechniques, 13:* 866-869, 1992.
26. Powell, S. M., Petersen, G. M., Krush, A. J., Booker, S., Jen, J., Giardiello, F. M., Hamilton, S. R., Vogelstein, B., and Kinzler, K. W. Molecular diagnosis of familial adenomatous polyposis. *New Engl. J. Med., 329:* 1982-1987, 1993.
27. Smith, K. J., Johnson, K. A., Bryan, T. M., Hill, D. E., Markowitz, S., Wilson, J. K. V., Paraskeva, C., Petersen, G. M., Hamilton, S. R., Vogelstein, B., and Kinzler, K. W. The *APC* gene product in normal and tumor cells. *Proc. Natl. Acad. Sci. USA, 90:* 2846-2850, 1993.
28. Pinkel, D., Landegent, J., Collins, C., Fuscoe, J., Segraves, R., Lucas, J., and Gray, J. Fluorecence in situ hybridization with human chromosome-specific libraries: Detection of trisomy 21 and translocation of chromosome 4. *Proc. Natl. Acad. Sci. USA, 85:* 9138-9142, 1988.
29. Meltzer, P. S., Guan, X.-Y., Burgess, A., and Trent, J. M. Micro-FISH: a novel stategy to identify cryptic chromosomal rearrangements. *Nature Genet., 1:* 24-28, 1992.

30. Altschul, S. F., Gish, W., Miller, W., Myers, E.W., and Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol. 215:* 403-410, 1990.

31. Schuler, G. D., Altschul, S. F., and Lipman D. J. A workbench for multiple alignment construction and analysis. *Proteins Struct. Funct. Genet. 9:* 180-190, 1991.

32. Fattaey, A. R., Helin, K., Dembski, M. S., Dyson, N., Harlow, E., Vuosolo, G. A., Hanobik, M. G., Haskell, K. M., Oliff, A., Defeo-Jones, D., and Jones, R. E. Characterization of the retinoblastoma binding proteins RBP1 and RBP2. *Oncogene, 8:* 3149-3156, 1993.

33. Kim, Y. W., Otterson, G. A., Kratzke, R. A., Coxon, A. B., and Kaye, F. J. Differential specificity for binding of retinoblastoma binding protein 2 to RB, p107, and TATA-binding protein. *Mol. Cell. Biol., 14:* 7256-7264, 1994.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: The Johns Hopkins University

(ii) TITLE OF INVENTION: EB1 Gene Product Binds to APC

(iii) NUMBER OF SEQUENCES: 12

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Banner & Allegretti, Ltd.
(B) STREET: 1001 G Street, N.W.
(C) CITY: Washington
(D) STATE: D.C.
(E) COUNTRY: U.S.
(F) ZIP: 20001-4597

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE: 22-MAY-1996
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Kagan, Sarah A.
(B) REGISTRATION NUMBER: 32,141
(C) REFERENCE/DOCKET NUMBER: 01107.49255

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 202.508.9100
(B) TELEFAX: 202.508.9299

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2540 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(vii) IMMEDIATE SOURCE:
(B) CLONE: EB1

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: 20q11.2

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 65..868

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
ACGAGACGAA GACGGAACCG GAGCCGGTTG CGGGCAGTGG ACGCGGTTCT GCCGAGAGCC      60

GAAG ATG GCA GTG AAC GTA TAC TCA ACG TCA GTG ACC AGT GAT AAC CTA      109
     Met Ala Val Asn Val Tyr Ser Thr Ser Val Thr Ser Asp Asn Leu
       1               5                  10                  15

AGT CGA CAT GAC ATG CTG GCC TGG ATC AAT GAG TCT CTG CAG TTG AAT      157
Ser Arg His Asp Met Leu Ala Trp Ile Asn Glu Ser Leu Gln Leu Asn
                   20                  25                  30

CTG ACA AAG ATC GAA CAG TTG TGC TCA GGG GCT GCG TAT TGT CAG TTT      205
Leu Thr Lys Ile Glu Gln Leu Cys Ser Gly Ala Ala Tyr Cys Gln Phe
               35                  40                  45

ATG GAC ATG CTG TTC CCT GGC TCC ATT GCC TTG AAG AAA GTG AAA TTC      253
Met Asp Met Leu Phe Pro Gly Ser Ile Ala Leu Lys Lys Val Lys Phe
           50                  55                  60

CAA GCT AAG CTA GAA CAC GAG TAC ATC CAG AAC TTC AAA ATA CTA CAA      301
Gln Ala Lys Leu Glu His Glu Tyr Ile Gln Asn Phe Lys Ile Leu Gln
       65                  70                  75

GCA GGT TTT AAG AGA ATG GGT GTT GAC AAA ATA ATT CCT GTG GAC AAA      349
Ala Gly Phe Lys Arg Met Gly Val Asp Lys Ile Ile Pro Val Asp Lys
   80                  85                  90                  95

TTA GTA AAA GGA AAG TTT CAG GAC AAT TTT GAA TTC GTT CAG TGG TTC      397
Leu Val Lys Gly Lys Phe Gln Asp Asn Phe Glu Phe Val Gln Trp Phe
                  100                 105                 110

AAG AAG TTT TTC GAT GCA AAC TAT GAT GGA AAA GAC TAT GAC CCT GTG      445
Lys Lys Phe Phe Asp Ala Asn Tyr Asp Gly Lys Asp Tyr Asp Pro Val
              115                 120                 125

GCT GCC AGA CAA GGT CAA GAA ACT GCA GTG GCT CCT TCC CTT GTT GCT      493
Ala Ala Arg Gln Gly Gln Glu Thr Ala Val Ala Pro Ser Leu Val Ala
          130                 135                 140

CCA GCT CTG AAT AAA CCG AAG AAA CCT CTC ACT TCT AGC AGT GCA GCT      541
Pro Ala Leu Asn Lys Pro Lys Lys Pro Leu Thr Ser Ser Ser Ala Ala
      145                 150                 155

CCC CAG AGG CCC ATC TCA ACA CAG AGA ACC GCT GCG GCT CCT AAG GCT      589
Pro Gln Arg Pro Ile Ser Thr Gln Arg Thr Ala Ala Ala Pro Lys Ala
160                 165                 170                 175

GGC CCT GGT GTG GTG CGA AAG AAC CCT GGT GTG GGC AAC GGA GAC GAC      637
Gly Pro Gly Val Val Arg Lys Asn Pro Gly Val Gly Asn Gly Asp Asp
                  180                 185                 190

GAG GCA GCT GAG TTG ATG CAG CAG GTC AAC GTA TTG AAA CTT ACT GTT      685
Glu Ala Ala Glu Leu Met Gln Gln Val Asn Val Leu Lys Leu Thr Val
              195                 200                 205
```

```
GAA GAC TTG GAG AAA GAG AGG GAT TTC TAC TTC GGA AAG CTA CGG AAC          733
Glu Asp Leu Glu Lys Glu Arg Asp Phe Tyr Phe Gly Lys Leu Arg Asn
        210                 215                 220

ATT GAA TTG ATT TGC CAG GAG AAC GAG GGG GAA AAC GAC CCT GTA TTG          781
Ile Glu Leu Ile Cys Gln Glu Asn Glu Gly Glu Asn Asp Pro Val Leu
    225                 230                 235

CAG AGG ATT GTA GAC ATT CTG TAT GCC ACA GAT GAA GGC TTT GTG ATA          829
Gln Arg Ile Val Asp Ile Leu Tyr Ala Thr Asp Glu Gly Phe Val Ile
240                 245                 250                 255

CCT GAT GAA GGG GGC CCA CAG GAG GAG CAA GAA GAG TAT TAACAGCCTG           878
Pro Asp Glu Gly Gly Pro Gln Glu Glu Gln Glu Glu Tyr
                260                 265

GACCAGCAGA GCAACATCGG AATTCTTCAC TCCAAATCAT GTGCTTAACT GTAAAATACT        938

CCCTTTTGTT ATCCTTAGAG GACTCACTGG TTTCTTTTCA TAAGCAAAAA GTACCTCTTC        998

TTAAAGTGCA CTTTGCAGAC GTTTCACTCC TTTTCCAATA AGTTTGAGTT AGGAGCTTTT       1058

ACCTTGTAGC AGAGCAGTAT TAACATCTAG TTGGTTCACC TGGAAAACAG AGAGGCTGAC       1118

CGTGGGGCTC ACCATGCGGA TGCGGGTCAC ACTGAATGCT GGAGAGATGT ATGTAATATG       1178

CTGAGGTGGC GACCTCAGTG GAGAAATGTA AAGACTGAAT TGAATTTTAA GCTAATGTGA       1238

AATCAGAGAA TGTTGTAATA AGTAAATGCC TTAAGAGTAT TTAAAATATG CTTCCACATT       1298

TCAAAATATA AAATGTAACA TGACAAGAGA TTTTGCGTTT GACATTGTGT CTGGGAAGGA       1358

AGGGCCAGAC CTTGGAACCT TTGGAACCTG CTGTCAACAG GTCTTACAGG GCTGCTTGAA       1418

CCCTCATAGG CCTAGGCTTT GGTCTAAAAG GAACATTTAA AAAGTTGCCC TGTAAAGTTA       1478

TTTGGTGTCA TTGACCAATT GCATCCCAGC TAAAAAGCAA GAGGCATCGT TGCCTGGATA       1538

ATAGAGGATG TGTTTCAGCC CTGAGATGTT ACAGTTGAAG AGCTTGGTTT CATTGAGCAT       1598

TTCTCTATTT TTCCAGTTAT CCCGAAATTT CTATGTATTA TTTTTTGGGG AAGTGAGGTG       1658

TGCCCAGTTT TTTAATCTAA CAACTACTTT TGGGGACTTG CCCACATCTC TGGGATTTGA       1718

ATGGGGATTG TATCCCATTT TACTGTCTTT TAGGTTTACA TTTACCACGT TTCTCTTCTC       1778

TGCTCCCCTT GCCCACTGGG ACTCCTCTTT GGCTCCTTGA AGTTTGCTGC TTAGAGTTGG       1838

AAGTGCAGCA GGCAGGTGAT CATGCTGCAA GTTCTTTCTG GACCTCTGGC AAAGGGAGTG       1898

GTCAGTGAAG GCCATCGTTA CCTTGGGATC TGCCAGGCTG GGGTGTTTTC GGTATCTGCT       1958

GTTCACAGCT CTCCACTGTA ATCCGAATAC TTTGCCAGTG CACTAATCTC TTTGGAGATA       2018

AAATTCATTA GTGTGTTACT AAATGTTAAT TTTCTTTTGC GGAAAATACA GTACCGTGTC       2078

TGAATTAATT ATTAATATTT AAAATACTTC ATTCCTTAAC TCTCCCTCAT TTGCTTTGCC       2138

CACAGCCTAT TCAGTTCCTT TGTTTGGCAG GATTCTGCAA AATGTGTCTC ACCCACTACT       2198

GAGATTGTTC AGCCCCTGAT GTATTTGTAT TGATTTGTTT CTGGTGGTAG CTTGTCCTGA       2258
```

AATGTGTGTA GAAAGCAAGT ATTTTATGAT AAAAATGTTG TGTAGTGCAT GCTCTGTGTG 2318

GAATTCAGAG GAAAACCCAG ATTCAGTGAT TAACAATGCC AAAAAATGCA AGTAACTAGC 2378

CATTGTTCAA ATGACAGTGG TGCTATTTCT CTTTTGTGGC CTTTTAGACT TTTGTTGCCC 2438

TAAAATTCCA TTTTATTGGG AACCCATTTT CCACCTGGTC TTTCTTGACA GGGTTTTTTT 2498

CTACTTTAAA CAGTTTCTAA ATAAAATTCT GTATTTCAAA AA 2540

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 268 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ala Val Asn Val Tyr Ser Thr Ser Val Thr Ser Asp Asn Leu Ser
1 5 10 15

Arg His Asp Met Leu Ala Trp Ile Asn Glu Ser Leu Gln Leu Asn Leu
20 25 30

Thr Lys Ile Glu Gln Leu Cys Ser Gly Ala Ala Tyr Cys Gln Phe Met
35 40 45

Asp Met Leu Phe Pro Gly Ser Ile Ala Leu Lys Lys Val Lys Phe Gln
50 55 60

Ala Lys Leu Glu His Glu Tyr Ile Gln Asn Phe Lys Ile Leu Gln Ala
65 70 75 80

Gly Phe Lys Arg Met Gly Val Asp Lys Ile Ile Pro Val Asp Lys Leu
85 90 95

Val Lys Gly Lys Phe Gln Asp Asn Phe Glu Phe Val Gln Trp Phe Lys
100 105 110

Lys Phe Phe Asp Ala Asn Tyr Asp Gly Lys Asp Tyr Asp Pro Val Ala
115 120 125

Ala Arg Gln Gly Gln Glu Thr Ala Val Ala Pro Ser Leu Val Ala Pro
130 135 140

Ala Leu Asn Lys Pro Lys Lys Pro Leu Thr Ser Ser Ser Ala Ala Pro
145 150 155 160

Gln Arg Pro Ile Ser Thr Gln Arg Thr Ala Ala Ala Pro Lys Ala Gly
165 170 175

Pro Gly Val Val Arg Lys Asn Pro Gly Val Gly Asn Gly Asp Asp Glu
180 185 190

Ala Ala Glu Leu Met Gln Gln Val Asn Val Leu Lys Leu Thr Val Glu
195 200 205

```
Asp Leu Glu Lys Glu Arg Asp Phe Tyr Phe Gly Lys Leu Arg Asn Ile
    210                 215                 220

Glu Leu Ile Cys Gln Glu Asn Glu Gly Glu Asn Asp Pro Val Leu Gln
225                 230                 235                 240

Arg Ile Val Asp Ile Leu Tyr Ala Thr Asp Glu Gly Phe Val Ile Pro
                245                 250                 255

Asp Glu Gly Gly Pro Gln Glu Glu Gln Glu Glu Tyr
            260                 265
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 149 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(vii) IMMEDIATE SOURCE:
(B) CLONE: EB2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Ile Ala Trp Val Asn Asp Ile Val Ser Leu Asn Tyr Thr Lys Val Glu
1               5                   10                  15

Gln Leu Cys Ser Gly Ala Ala Tyr Cys Gln Phe Met Asp Met Leu Phe
            20                  25                  30

Pro Gly Cys Ile Ser Leu Lys Lys Val Lys Phe Gln Ala Lys Leu Glu
        35                  40                  45

His Glu Tyr Ile His Asn Phe Lys Leu Leu Gln Ala Ser Phe Lys Arg
    50                  55                  60

Met Asn Val Asp Lys Val Ile Pro Val Glu Lys Leu Val Lys Gly Arg
65                  70                  75                  80

Phe Gln Asp Asn Leu Asp Phe Ile Gln Trp Phe Lys Lys Phe Tyr Asp
                85                  90                  95

Ala Asn Tyr Asp Gly Lys Glu Tyr Asp Pro Val Glu Ala Arg Gln Gly
            100                 105                 110

Gln Asp Ala Ile Pro Pro Pro Asp Pro Gly Glu Gln Ile Phe Asn Leu
        115                 120                 125

Pro Lys Lys Ser His His Ala Asn Ser Pro Thr Ala Gly Ala Ala Lys
    130                 135                 140
```

Phe Lys Phe Gln Xaa
145

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 344 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Saccharomyces cerevisiae

(vii) IMMEDIATE SOURCE:
(B) CLONE: Yer016p

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Ser Ala Gly Ile Gly Glu Ser Arg Thr Glu Leu Leu Thr Trp Leu
1 5 10 15

Asn Gly Leu Leu Asn Leu Asn Tyr Lys Lys Ile Glu Glu Cys Gly Thr
20 25 30

Gly Ala Ala Tyr Cys Gln Ile Met Asp Ser Ile Tyr Gly Asp Leu Pro
35 40 45

Met Asn Arg Val Lys Phe Asn Ala Thr Ala Glu Tyr Glu Phe Gln Thr
50 55 60

Asn Tyr Lys Ile Leu Gln Ser Cys Phe Ser Arg His Gly Ile Glu Lys
65 70 75 80

Thr Val Tyr Val Asp Lys Leu Ile Arg Cys Lys Phe Gln Asp Asn Leu
85 90 95

Glu Phe Leu Gln Trp Leu Lys Lys His Trp Ile Arg His Lys Asp Glu
100 105 110

Ser Val Tyr Asp Pro Asp Ala Arg Arg Lys Tyr Arg Pro Ile Ile Thr
115 120 125

Asn Asn Ser Ala Thr Lys Pro Arg Thr Val Ser Asn Pro Thr Thr Ala
130 135 140

Lys Arg Ser Ser Ser Thr Gly Thr Gly Ser Ala Met Ser Gly Gly Leu
145 150 155 160

Ala Thr Arg His Ser Ser Leu Gly Ile Asn Gly Ser Arg Lys Thr Ser
165 170 175

Val Thr Gln Gly Gln Leu Val Ala Ile Gln Ala Glu Leu Thr Lys Ser
180 185 190

```
Gln Glu Thr Ile Gly Ser Leu Asn Glu Glu Ile Glu Gln Tyr Lys Gly
        195                 200                 205

Thr Val Ser Thr Leu Glu Ile Glu Arg Glu Phe Tyr Phe Asn Lys Leu
    210                 215                 220

Arg Asp Ile Glu Ile Leu Val His Thr Thr Gln Asp Leu Ile Asn Glu
225                 230                 235                 240

Gly Val Tyr Lys Phe Asn Asp Glu Thr Ile Thr Gly His Gly Asn Gly
                245                 250                 255

Asn Gly Gly Ala Leu Leu Arg Phe Val Lys Lys Val Glu Ser Ile Leu
            260                 265                 270

Tyr Ala Thr Ala Glu Gly Phe Glu Met Asn Asp Gly Glu Asp Glu Leu
        275                 280                 285

Asn Asp Lys Asn Leu Gly Glu His Gly Thr Val Pro Asn Gln Gly Gly
    290                 295                 300

Tyr Ala Asn Ser Asn Gly Glu Val Asn Gly Asn Glu Gly Ser Asn His
305                 310                 315                 320

Asp Val Ile Met Gln Asn Asp Glu Gly Glu Val Gly Val Ser Asn Asn
                325                 330                 335

Leu Ile Ile Asp Glu Glu Thr Phe
            340
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 112 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(vii) IMMEDIATE SOURCE:
(B) CLONE: z19434

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
Asp Glu Asp Pro Pro Pro Arg Ser Arg Arg Pro Glu Pro Gln Pro Leu
1               5                   10                  15

Pro Gln Arg Pro Arg His Leu Ser Pro Pro Pro Pro Pro Pro Pro Glu
            20                  25                  30

Pro Pro Arg Ala Leu Trp Gly Met Ala Val Asn Val Tyr Ser Thr Ser
        35                  40                  45
```

```
Val Thr Ser Glu Asn Leu Ser Arg His Asp Met Leu Ala Trp Val Asn
    50                  55                  60

Asp Ser Leu His Leu Asn Tyr Thr Lys Ile Glu Gln Leu Cys Ser Gly
65                  70                  75                  80

Ala Ala Tyr Cys Gln Phe Met Asp Met Leu Phe Pro Gly Cys Val His
                85                  90                  95

Leu Arg Lys Val Lys Phe Gln Gly Lys Leu Gly His Xaa Tyr Ile His
            100                 105                 110
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 120 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: YES

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo sapiens

(vii) IMMEDIATE SOURCE:
(B) CLONE: M85402

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Asn Phe Lys Val Leu Gln Xaa Ala Phe Lys Lys Met Gly Val Asp Lys
1               5                   10                  15

Ile Ile Pro Val Glu Lys Leu Val Lys Gly Lys Phe Gln Asp Asn Phe
            20                  25                  30

Xaa Phe Ile Gln Trp Phe Lys Lys Xaa Phe Asp Ala Asn Tyr Asp Gly
        35                  40                  45

Lys Asp Tyr Asn Pro Leu Leu Ala Arg Gln Gly Gln Asp Val Ala Pro
    50                  55                  60

Pro Pro Asn Pro Val Pro Gln Arg Thr Ser Pro Thr Gly Pro Lys Asn
65                  70                  75                  80

Met Gln Thr Ser Gly Arg Leu Ser Asn Val Ala Pro Pro Cys Ile Leu
                85                  90                  95

Arg Lys Xaa Pro Pro Ser Ala Arg Asn Gly Gly His Glu Thr Cys Pro
            100                 105                 110

Asn Ser Leu Asn Ser Asn Gln Gln
        115                 120
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 54 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGATCCTAAT ACGACTCACT ATAGGGAGAC CACCATGGCA GTGAACGTAT ACTC 54

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ATTTCTCCAC TGAGGTCGC 19

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CGAGATCTAA GATGGCAGTG AACGTATA 28

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GCAGATCTTT AATACTCTTC TTGATCCTCC 30

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AAAACAGAGA GGCTGACCG 19

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

ATTTCTCCAC TGAGGTCGC 19

CLAIMS

1. A nucleic acid molecule comprising an *EB1* DNA according to SEQ ID NO:1.

2. The nucleic acid molecule of claim 1 further comprising a vector containing an origin of replication.

3. An isolated and purified nucleic acid molecule comprising at least 18 contiguous nucleotides of *EB1* coding sequence according to SEQ ID NO:1.

4. An isolated and purified nucleic acid molecule which comprises a coding sequence which encodes at least 20 contiguous amino acids of *EB1* according to SEQ ID NO:2.

5. An isolated and purified EB1 protein according to SEQ ID NO:2.

6. An isolated and purified EB1 polypeptide comprising at least 20 contiguous amino acids according to SEQ ID NO:2.

7. An isolated and purified nucleic acid molecule consisting at least 12 contiguous nucleotides of *EB1* coding sequence according to SEQ ID NO:1.

8. An isolated and purified nucleic acid molecule consisting of a coding sequence for at least 6 contiguous amino acids of *EB1* according to SEQ ID NO:2.

9. An isolated and purified EB1 polypeptide consisting of at least 6 contiguous amino acids according to SEQ ID NO:2.

10. A method for determining a predisposition to neoplasms, comprising the step of:

determining a mutation in an *EB1* allele of a human tissue, wherein wild-type *EB1* is as shown in SEQ ID NO:1.

11. The method of claim 10 wherein said human tissue is selected from the group consisting of blood, chorionic villi, fetal trophoblasts, amniotic fluid, and a blastomere of a pre-implantation embryo.

12. A method for determining a predisposition to neoplasms, comprising the step of:

assaying protein complexes in a cell, wherein said protein complexes comprise APC and EB1, wherein reduction of said complexes in the

cell relative to a cell which contains wild-type *APC* and wild-type *EB1* alleles indicates a predisposition to neoplasms.

13. The method of claim 12 wherein said step of assaying is performed by immunoprecipitation followed by immunoblotting.

14. The method of claim 13 wherein immunoprecipitation is performed with an antibody specifically immunoreactive with APC.

15. The method of claim 13 wherein immunoprecipitation is performed with an antibody specifically immunoreactive with EB1.

16. The method of claim 14 wherein immunoblotting is performed with an antibody specifically immunoreactive with EB1.

17. The method of claim 15 wherein immunoblotting is performed with an antibody specifically immunoreactive with APC.

18. An antibody preparation which is specifically immunoreactive with EB1.

19. The preparation of claim 17 wherein said antibody is monoclonal.

20. The preparation of claim 17 wherein said antibody is polyclonal.

21. A method for determining a predisposition to cancer, comprising the steps of:

testing a human tissue to determine if the tissue expresses less EB1 gene products than a normal human tissue.

22. The method of claim 21 wherein the step of testing utilizes an antibody which is specifically immunoreactive with EB1 protein.

23. The method of claim 21 wherein the step of testing utilizes a nucleic acid probe which specifically hybridizes to an *EB1* mRNA, said probe having a sequence of at least 12 contiguous nucleotides selected from SEQ ID NO:1.

24. A method for diagnosing a neoplasm, comprising the step of:

determining mutations in *EB1* alleles of a human tissue, wherein wild-type *EB1* is as shown in SEQ ID NO:1.

25. A method for diagnosing a neoplasm, comprising the step of:

assaying protein complexes in a cell, wherein said protein complexes comprise APC and EB1, wherein absence of said complexes indicates a neoplasm.

26. The method of claim 25 wherein said step of assaying is performed by immunoprecipitation followed by immunoblotting.

27. The method of claim 26 wherein immunoprecipitation is performed with an antibody specifically immunoreactive with APC.

28. The method of claim 26 wherein immunoprecipitation is performed with an antibody specifically immunoreactive with EB1.

29. The method of claim 27 wherein immunoblotting is performed with an antibody specifically immunoreactive with EB1.

30. The method of claim 28 wherein immunoblotting is performed with an antibody specifically immunoreactive with APC.

31. A method for diagnosing a neoplasm, comprising the step of:

testing a human tissue to determine if the tissue expresses EB1 gene products, wherein a tissue which expresses no EB1 gene products is neoplastic.

32. The method of claim 31 wherein the step of testing utilizes an antibody which is specifically immunoreactive with EB1 protein.

33. The method of claim 31 wherein the step of testing utilizes a nucleic acid probe which specifically hybridizes to an *EB1* mRNA, said probe having a sequence of at least 12 contiguous nucleotides selected from SEQ ID NO:1.

34. A method to assess treatment options for a cancer, comprising the step of:

contacting a lysate of cancer cells with EB1 protein and detecting the formation of protein complexes comprising said EB1 protein, a lysate which fails to form complexes indicating cancer cells which are good candidates for treatment with cyclooxygenase inhibitors.

35. A method to assess the status of *APC* alleles in a cell, comprising the step of:

contacting a lysate of cells with EB1 protein and detecting the formation of protein complexes comprising said EB1 protein, a lysate which fails to form complexes indicating cancer cells which may lack wild-type *APC*.

```
   1                                                                                        Met Ala Val Asn Val Tyr     6
   1  A CGA GAC GAA GAC GGA ACC GGA GCC GGT TGC GGG CAG TGG ACG CGG TTC TGC CGA GAG CCG AAG ATG GCA GTG AAC GTA TAC    82

   7 Ser Thr Ser Val Thr Ser Asp Asn Leu Ser Arg His Asp Met Leu Ala Trp Ile Asn Glu Ser Leu Gln Leu Asn Leu Thr Lys    34
  83 TCA ACG TCA GTG ACC AGT GAT AAC CTA AGT CGA CAT GAC ATG CTG GCC TGG ATC AAT GAG TCT CTG CAG TTG AAT CTG ACA AAG   166

  35 Ile Glu Gln Leu Cys Ser Gly Ala Ala Tyr Cys Gln Phe Met Asp Met Leu Phe Pro Gly Ser Ile Ala Leu Lys Lys Val Lys    62
 167 ATC GAA CAG TTG TGC TCA GGG GCT GCG TAT TGT CAG TTT ATG GAC ATG CTG TTC CCT GGC TCC ATT GCC TTG AAG AAA GTG AAA   250

  63 Phe Gln Ala Lys Leu Glu His Glu Tyr Ile Gln Asn Phe Lys Ile Leu Gln Ala Gly Phe Lys Arg Met Gly Val Asp Lys Ile    90
 251 TTC CAA GCT AAG CTA GAA CAC GAG TAC ATC CAG AAC TTC AAA ATA CTA CAA GCA GGT TTT AAG AGA ATG GGT GTT GAC AAA ATA   334

  91 Ile Pro Val Asp Lys Leu Val Lys Gly Lys Phe Gln Asp Asn Phe Glu Phe Val Gln Trp Phe Lys Lys Phe Phe Asp Ala Asn   118
 335 ATT CCT GTG GAC AAA TTA GTA AAA GGA AAG TTT CAG GAC AAT TTT GAA TTC GTT CAG TGG TTC AAG AAG TTT TTC GAT GCA AAC   418

 119 Tyr Asp Gly Lys Asp Tyr Asp Pro Val Ala Ala Arg Gln Gly Gln Glu Thr Ala Val Ala Pro Ser Leu Val Ala Pro Ala Leu   146
 419 TAT GAT GGA AAA GAC TAT GAC CCT GTG GCT GCC AGA CAA GGT CAA GAA ACT GCA GTG GCT CCT TCC CTT GTT GCT CCA GCT CTG   502

 147 Asn Lys Pro Lys Lys Pro Leu Thr Ser Ser Ser Ala Ala Pro Gln Arg Pro Ile Ser Thr Gln Arg Thr Ala Ala Ala Pro Lys   174
 503 AAT AAA CCG AAG AAA CCT CTC ACT TCT AGC AGT GCA GCT CCC CAG AGG CCC ATC TCA ACA CAG AGA ACC GCT GCG GCT CCT AAG   586

 175 Ala Gly Pro Gly Val Val Arg Lys Asn Pro Gly Val Gly Asn Gly Asp Asp Glu Ala Ala Glu Leu Met Gln Gln Val Asn Val   202
 587 GCT GGC CCT GGT GTG GTG CGA AAG AAC CCT GGT GTG GGC AAC GGA GAC GAC GAG GCA GCT GAG TTG ATG CAG CAG GTC AAC GTA   670

 203 Leu Lys Leu Thr Val Glu Asp Leu Glu Lys Glu Arg Asp Phe Tyr Phe Gly Lys Leu Arg Asn Ile Glu Leu Ile Cys Gln Glu   230
 671 TTG AAA CTT ACT GTT GAA GAC TTG GAG AAA GAG AGG GAT TTC TAC TTC GGA AAG CTA CGG AAC ATT GAA TTG ATT TGC CAG GAG   754

 231 Asn Glu Gly Glu Asn Asp Pro Val Leu Gln Arg Ile Val Asp Ile Leu Tyr Ala Thr Asp Glu Gly Phe Val Ile Pro Asp Glu   258
 755 AAC GAG GGG GAA AAC GAC CCT GTA TTG CAG AGG ATT GTA GAC ATT CTG TAT GCC ACA GAT GAA GGC TTT GTG ATA CCT GAT GAA   838

 259 Gly Gly Pro Gln Glu Glu Gln Glu Glu Tyr                                                                            268
 839 GGG GGC CCA CAG GAG GAG CAA GAA GAG TAT TAA CAG CCT GGA CCA GCA GAG CAA CAT CGG AAT TCT TCA CTC CAA ATC ATG TGC   922
 923 TTA ACT GTA AAA TAC TCC CTT TTG TTA TCC TTA GAG GAC TCA CTG GTT TCT TTT CAT AAG CAA AAA GTA CCT CTT CTT AAA GTG  1006
1007 CAC TTT GCA GAC GTT TCA CTC CTT TTC CAA TAA GTT TGA GTT AGG AGC TTT TAC CTT GTA GCA GAG CAG TAT TAA CAT CTA GTT  1090
1091 GGT TCA CCT GGA AAA CAG AGA GGC TGA CCG TGG GGC TCA CCA TGC GGA TGC GGG TCA CAC TGA ATG CTG GAG AGA TGT ATG TAA  1174
1175 TAT GCT GAG GTG GCG ACC TCA GTG GAG AAA TGT AAA GAC TGA ATT GAA TTT TAA GCT AAT GTG AAA TCA GAG AAT GTT GTA ATA  1258
1259 AGT AAA TGC CTT AAG AGT ATT TAA AAT ATG CTT CCA CAT TTC AAA ATA TAA AAT GTA ACA TGA CAA GAG ATT TTG CGT TTG ACA  1342
1343 TTG TGT CTG GGA AGG AAG GGC CAG ACC TTG GAA CCT TTG GAA CCT GCT GTC AAC AGG TCT TAC AGG GCT GCT TGA ACC CTC ATA  1426
1427 GGC CTA GGC TTT GGT CTA AAA GGA ACA TTT AAA AAG TTG CCC TGT AAA GTT ATT TGG TGT CAT TGA CCA ATT GCA TCC CAG CTA  1510
1511 AAA AGC AAG AGG CAT CGT TGC CTG GAT AAT AGA GGA TGT GTT TCA GCC CTG AGA TGT TAC AGT TGA AGA GCT TGG TTT CAT TGA  1594
1595 GCA TTT CTC TAT TTT TCC AGT TAT CCC GAA ATT TCT ATG TAT TAT TTT TTG GGG AAG TGA GGT GTG CCC AGT TTT TTA ATC TAA  1678
1679 CAA CTA CTT TTG GGG ACT TGC CCA CAT CTC TGG GAT TTG AAT GGG GAT TGT ATC CCA TTT TAC TGT CTT TTA GGT TTA CAT TTA  1762
1763 CCA CGT TTC TCT TCT CTG CTC CCC TTG CCC ACT GGG ACT CCT CTT TGG CTC CTT GAA GTT TGC TGC TTA GAG TTG GAA GTG CAG  1846
1847 CAG GCA GGT GAT CAT GCT GCA AGT TCT TTC TGG ACC TCT GGC AAA GGG AGT GGT CAG TGA AGG CCA TCG TTA CCT TGG GAT CTG  1930
1931 CCA GGC TGG GGT GTT TTC GGT ATC TGC TGT TCA CAG CTC TCC ACT GTA ATC CGA ATA CTT TGC CAG TGC ACT AAT CTC TTT GGA  2014
2015 GAT AAA ATT CAT TAG TGT GTT ACT AAA TGT TAA TTT TCT TTT GCG GAA AAT ACA GTA CCG TGT CTG AAT TAA TTA TTA ATA TTT  2098
2099 AAA ATA CTT CAT TCC TTA ACT CTC CCT CAT TTG CTT TGC CCA CAG CCT ATT CAG TTC CTT TGT TTG GCA GGA TTC TGC AAA ATG  2182
2183 TGT CTC ACC CAC TAC TGA GAT TGT TCA GCC CCT GAT GTA TTT GTA TTG ATT TGT TTC TGG TGG TAG CTT GTC CTG AAA TGT GTG  2266
2267 TAG AAA GCA AGT ATT TTA TGA TAA AAA TGT TGT GTA GTG CAT GCT CTG TGT GGA ATT CAG AGG AAA ACC CAG ATT CAG TGA TTA  2350
2351 ACA ATG CCA AAA AAT GCA AGT AAC TAG CCA TTG TTC AAA TGA CAG TGG TGC TAT TTC TCT TTT GTG GCC TTT TAG ACT TTT GTT  2434
2435 GCC CTA AAA TTC CAT TTT ATT GGG AAC CCA TTT TCC ACC TGG TCT TTC TTG ACA GGG TTT TTT TCT ACT TTA AAC AGT TTC TAA  2518
2519 ATA AAA [illegible] TGT ATT TCA AAA A                                                                             2540
```

FIGURE 1

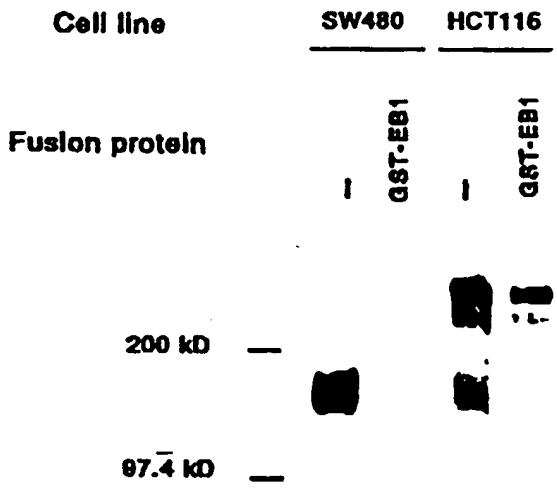

Cell line
SW480
HCT116
Fusion protein
-
GST-EB1
-
GST-EB1
200 kD
97.4 kD

FIGURE 2A

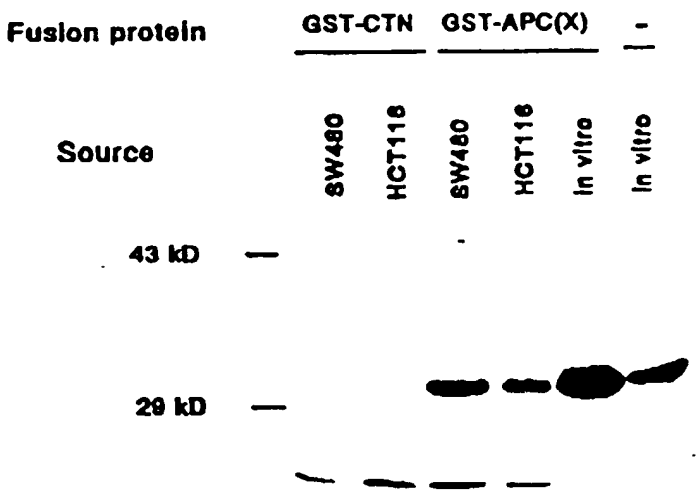

Fusion protein
GST-CTN
GST-APC(X)
-
Source
SW480
HCT116
SW480
HCT116
In vitro
In vitro
43 kD
29 kD

FIGURE 2B

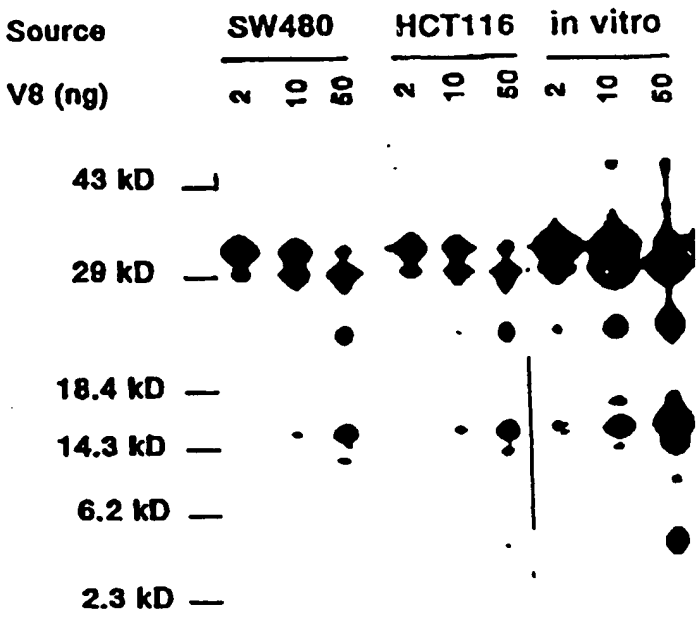

Source
SW480
HCT116
in vitro
V8 (ng)
2
10
50
2
10
50
2
10
50
43 kD
29 kD
18.4 kD
14.3 kD
6.2 kD
2.3 kD

FIGURE 2C

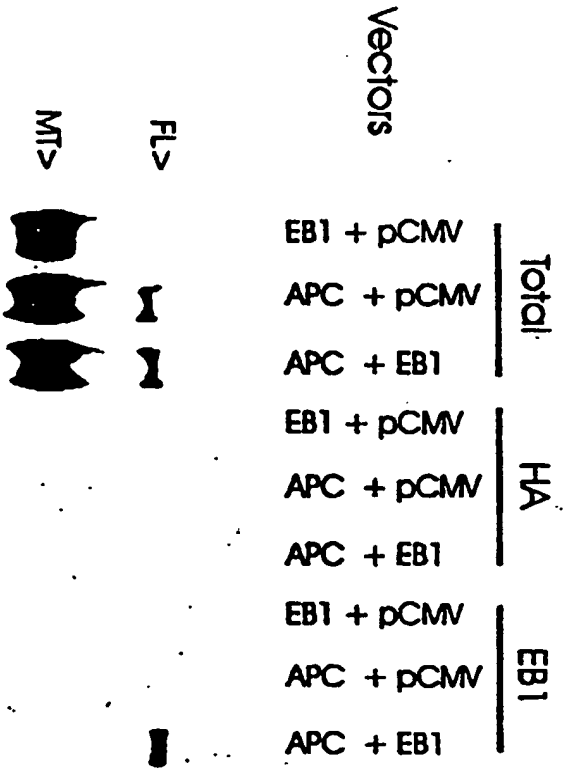

Vectors
MT>
FL>
EB1 + pCMV
APC + pCMV
APC + EB1
Total
EB1 + pCMV
APC + pCMV
APC + EB1
HA
EB1 + pCMV
APC + pCMV
APC + EB1
EB1

FIGURE 3

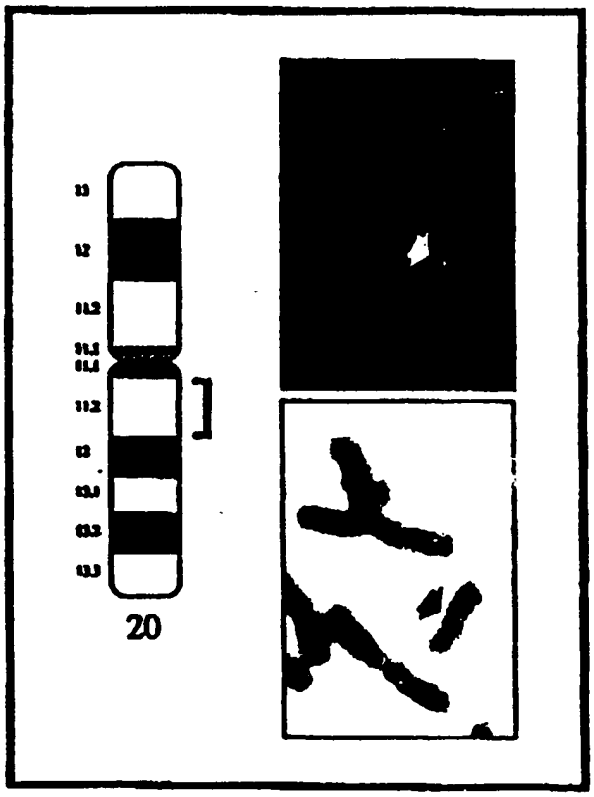


FIGURE 4

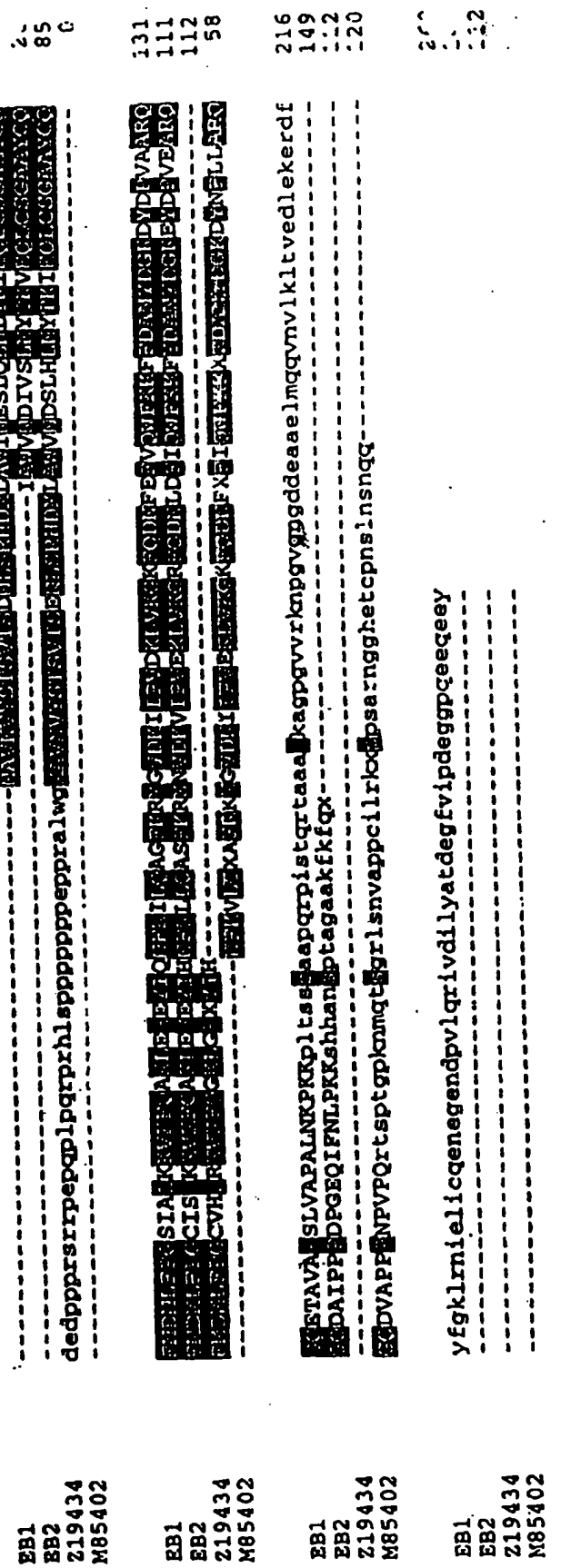

EB1
EB2
Z19434
M85402
dedpppsrrrpepqplpqrprhlsppppppeppralwg
yfgklrnielicqenegendpvlqrivdllyatdegfvipdeggpqeeqeey

FIGURE 5A

EB-1
Yer016p
EB-1
Yer016p
EB-1
Yer016p
EB-1
Yer016p
EB-1
Yer016p
85
76
161
144
226
229
268
314
26
344

FIGURE 5 B

INTERNATIONAL SEARCH REPORT

Inter ional Application No
PCT/US 96/07747

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/435 G01N33/574 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K G01N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EMBL DATABASE, Accession number T03463 Sequence reference HST03463 from human cDNA clone IB327; 24 August 1993 Compare nucleotides 0-415 with nucleotides XP002012454 2539-2125 of SEQ. ID. NO. 1 ---	3,4
X	EMBL DATABASE, Accession number D12076 Sequence reference HS000S163 from human HepG2 clone S163; 18 November 1992 Compare nucleotides 0-610 with nucleotides XP002012455 1925 to 2535 of SEQ.ID.NO.1 --- -/--	3,4

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
3 September 1996	16. 09. 96

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Cupido, M

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/US 96/07747

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE SWISSPROT,Accession number S50474 Sequence reference S50428 of S.cerevisiae cosmid 9537, identical to Yer016W;28-05-93 XP002012456 Compare amino acids 33-38 with amino acids 41-46 of SEQ. ID. NO.2. ---	8
P,X	CANCER RESEARCH, vol. 55, 15 July 1995, MD US, pages 2972-2977, XP002012453 L-K SU ET AL.: "APC binds to the novel protein EB1" see the whole document ---	1-35
A	WO,A,94 21814 (THE JOHN HOPKINS UNIVERSITY) 29 September 1994 see the whole document -----	1-35

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US96/07747

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.: 11
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claim 11 may be directed to a diagnostic method practised on the human/animal body the search has been carried out and based on the alleged effects of the composition.

2. [ ] Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. [ ] As all searchable claims could be searches without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

...formation on patent family members

Inter ...nal Application No
PCT/US 96/07747

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9421814	29-09-94	CA-A- 2158429	29-09-94
		EP-A- 0689606	03-01-96